

Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики

Моя гипотеза состоит в том, что слово «предисловие» является лингвоспецифичным, а слово «шкаф» - нет.

Я выбрала лингвоспецифичное слово «предисловие» для своего исследования. Количество вхождений этого слова равно 36. В таблице ниже мы можем наблюдать различные модели перевода этого слова на английский язык, а также частоту встречаемости по данным НКРЯ.

Модель перевода	Частота употребляемости
Preface	9
Editor`s note	2
introduction	6
foreword	15
foreshadowed	1
preamble	1
preliminary account	1
prefatory remarks	1
entry	1

Таблица с формулами и полученными результатами:

формулы	результаты
$(F (M_{max}))/(\text{Num}M) 15/5$	3
$(F (O))/\text{Num}M$, где $F (O) 36/5$	7,2
$(F (M_{max})/F (M_{sec})) 15/9$	1,6
$(F (M_{max})/F (O)) 15/36$	0,41

Неспецифичное слово, которое я выбрала, — это «шкаф». Количество вхождений в корпусе составляет 168 (не включая слова, приписанные автором-переводчиком). В таблице ниже мы можем наблюдать различные модели перевода этого слова на английский язык, а также частоту встречаемости по данным НКРЯ.

Модель перевода	Частота употребляемости
Closet/closetful	45
Delivery men	1
drawer	8
wardrobe	19
cupboard	30
bookcase	18
shelf	7
cabinet	15
armoire	2
clothes-chest	1

dresser	6
case	2
press	11
benches	1
repository	1
chest	1

формулы	результаты
$(F(M_{max}))/(\text{Num}M)$ 45/11	4,09
$(F(O))/\text{Num}M$, где $F(O)$ 168/11	15,27
$(F(M_{max}))/F(M_{sec})$ 45/30	1,5
$(F(M_{max}))/F(O)$ 45/168	0,27

Выводы: В ходе полученных результатов, можно сказать, что гипотеза об определении лингвоспецифичности подобранных мною слов подтвердилась не полностью. Слово «предисловие» нельзя назвать полностью лингвоспецифичным, так как самая частотная модель перевода - «foreword» составляет около 42%, что немного противоречит изначальному заданию («Предполагается, что у «лингвоспецифичного» слова должен наблюдаться большой список (список оказался небольшим) возможных моделей перевода, и на каждую будет в среднем приходится сравнительно немного контекстов (а здесь действительно так, так как больше половины моделей употреблялись всего лишь по одному разу), а самая частотная из всех моделей будет занимать небольшой процент (42% - это довольно много) от общего числа возможных соответствий, ее частота не будет сильно отличаться от частот остальных» (отличается довольно сильно)), в отличие от слова «шкаф», где самая частотная модель перевода – «closet» употребляется только в 26,8%. Также, оказалось, что у слова «шкаф» примерно в 2 раза больше моделей перевода, чем у слова «предисловие», соответственно, можно считать его лингвоспецифичным.